ChinaXiv合作期刊 图书情报工作
第65卷 第3期 2021年2月

政府数据开放准备度关键影响因素识别*

——以省级地方政府为例

■ 朱玲玲[1] 茆意宏[1] 朱永凤[2] 马坤坤[1]

[1] 南京农业大学信息管理学院 南京 210095 [2]南京航空航天大学经济与管理学院 南京 211106

摘 要：[目的/意义]政府数据开放的准备度是为向社会提供更高价值数据所做准备的完善程度，考察影响地方政府数据开放准备度的差异影响因素并据此提出推进路径，旨在促进数据开放实践可持续。[方法/过程]运用定性比较分析方法，以13个省级政府数据开放准备度政策等作为研究样本，分析影响政府数据开放准备度的多种因素的"联合效应"和各因素之间的"互动关系"。[结果/结论]研究发现，安全保障是政府数据开放准备度的必要条件，数据管理、数据利用、部门职责、安全保障4种要素需同时具备，文本效力、领导公开支持以辅助条件形式与数据管理、数据利用、安全保障、部门职责要素组合在一起使用。

关键词：政府数据开放 准备度 QCA 影响因素
分类号：G20
DOI：10.13266/j.issn.0252-3116.2021.03.010

政府数据开放是将由政府生成结构化、机器可读的数据集提供给利益相关者下载获取、利用，以实现社会效益和经济效益[1]。2009年，全球开始开展政府数据开放运动，英国、美国、加拿大等国家积极出台相关政策措施推动政府数据开放实践。如何将政府数据有效地向社会公众开放也成为世界各国的发展战略问题。2015年8月，我国国务院印发《促进大数据发展行动纲要》，明确指出大数据成为提升政府治理能力的新途径，运用大数据技术推动政府数据开放共享，挖掘和实现数据资源的潜在价值[2]。2019年在全国两会上，政府数据开放被不断提及，政府数据资源共享开放已被列入我国大数据发展的"十大关键工程"[3]。伴随着以"互联网+"为代表的物联网、大数据、区块链、人工智能等新一代信息技术的兴起，政府数据开放的广度和深度需要进一步深化。由万维网基金会(The Word Wide Web Foundation)和开放数据研究院(Open Data Institute)联合发布的《开放数据晴雨表(Open Data Barometer)》评估框架[4]主要从准备度、执行力、影响力3个维度分析各国的开放数据活动和开放数据的全球趋势，明确使用"开放数据准备度"的指标衡量政府为实现开放数据所作的前期准备工作[5]，包括是否制定完备的政策与协议以确保开放数据长期可获得、是否为各级政府利用开放数据打好基础、公民及民间团体是否利用开放数据参与政策制定、企业及企业家是否充分利用开放数据带来的经济机遇4个方面的内容；准备度、执行力、影响力是政府数据开放发展的不同阶段，各维度之间存在关联性。简单来说，前期充分的准备度保障中期的执行力实施，高度的执行力能提高后期政府数据开放影响力，后期影响力的提升进一步促进准备度与执行力[6]。准备度是政府开放数据的基础和先决条件，是衡量政府为开放数据所采取的举措与付诸实施的政策，某个地方的准备度越高，相应的数据层、平台层与利用层的效果更好。目前探索政府数据开放准备度的研究成果不多，在有限的文献中，有代表性的政府数据开放准备度评估框架包括：世界银行的"开放数据准备度"评估框架(Open Data Readiness Assessment)[7]将各国政府、地区级、大城市政府等的开放数据项目作为评估对象，聚焦于开放数据门户网站及动态的开放数据生态系统；中国开放数林指数(China Open Data Index)是由复旦大学与国家信息中心

* 本文系江苏省研究生科研与实践创新计划项目"高校图书馆文化对服务绩效的作用机制研究"(项目编号：KYCX20_0617)研究成果之一。
作者简介：朱玲玲(ORCID：0000-0002-7709-4303)，博士研究生；茆意宏(ORCID：0000-0002-1596-8837)，教授，博士生导师，通讯作者，E-mail：maoyh@njau.edu.cn；朱永凤(ORCID：0000-0003-4684-2935)，博士研究生；马坤坤(ORCID：0000-0003-3256-9253)，博士。
收稿日期：2020-08-23 修回日期：2020-11-30 本文起止页码：75-83 本文责任编辑：徐健

数字中国研究院联合发布,认为政府数据开放准备度主要包括法规与政策、组织与执行、标准规范等方面[8]。我国地方政府也在积极开展数据开放实践探索,比如贵州、上海等地出台了专门针对数据开放的政策,广东、山东等省制定出台了政府数据开放的地方标准。

总体而言,我国的政府数据开放尚处于初步发展阶段,这一阶段准备度显得尤为重要,目前的研究仅限于从整体评估政府数据开放的准备度,还未讨论哪些要素的组合能够更有效地促进政府数据开放的准备工作。笔者拟以此为切入点,将文本效力、数据管理、数据利用、安全保障等 8 种政府数据开放准备度影响因素作为前因条件变量,将 13 个省级地方政府数据开放准备度指数作为结果变量,借助定性比较分析(Qualitative Comparative Analysis, QCA)方法从多原因角度分析条件变量与结果变量之间的关系优势,通过对前因条件变量组合进行比较分析,发现提高地方政府数据开放准备度的关键要素或者要素组合,进而探究改善我国政府数据开放准备度基础,以及政府数据开放准备工作的提升路径。

1 研究方法、研究样本选取及研究变量确定

1.1 研究方法

1987 年, R. Charles 首次提出定性比较分析方法(QCA)[9],以"与(Logical AND)""或(Logical OR)""非(Negation)"的集合运算逻辑为理论基础[10],通过比较分析各个案例中的多个因果条件,结合案例导向方法,寻求被解释变量和解释变量间的多重并发因果关系与多种组态,旨在寻找并解释导致某一结果发生或者不发生的前因条件及其组合。清晰集定性比较分析法(csQCA)是 R. Clarles 和 D. Kriss 开发的第一种 QCA 技术[11],本质上是以一种"以逻辑和整体方式简化复杂数据结构"的工具。csQCA 需要将条件变量与结果变量转换为二分变量,取值是 0 和 1。到目前为止,csQCA 是最广泛使用的 QCA 技术[12]。邓胜利等认为 QCA 方法是兼具定量与定性分析优势的实证研究方法,将其引入国内图书情报领域,并运用 QCA 方法分析网络社区健康信息搜寻影响因素问题[13];杨金龙等研究移动学习实现"采纳转化为持续"的动因及分析变量之间的非对称多重并发组态效应问题[14];杨洋洋等通过选取网络舆论事件,运用 QCA 方法探索网络舆情热度的关键因素及变量组合[15]。

笔者采用 csQCA 的主要原因如下:①传统实证方法对样本数量要求较大,且只能验证单个变量与因变量之间的关系,即只能处理 A→B,则 ~A→~B 的对称关系的问题,如果运用实证的研究方法无法得出影响政府数据开放准备度各因素之间的交互关系;②QCA 方法样本数量要求是 10 - 60,适用于中小数量的案例分析,以 13 个省级政府数据开放的准备度为研究样本,在样本数量上符合该方法;③此方法采用布尔代数运算方法,可以处理非对称因果关系,并通过对一致性(consistency)和覆盖度(coverage)两个参数得出最具解释力的前因条件构型,能够解析条件变量与结果变量之间复杂的多元、非线性因果关系。促使政府数据开放高准备度的因素是多重并发的,利用此方法,有利于识别并解释不同地区的政府数据开放准备度的相对有效因素与无效要素,进而总结出改善我国政府数据开放准备度基础的条件组合。

1.2 研究样本选取

2015 年 9 月国务院印发《促进大数据发展行动纲要》,我国正式启动政府数据开放行动,并取得了积极进展。由于部分省份、自治区、直辖市 2015 - 2019 年间未出台有关政府数据开放的针对性政策,结合《2019 年中国地方政府数据开放报告》中的准备度指数排名情况,本研究以上海、浙江、贵州、北京、福建等 13 个省级行政区地方为例,对它们出台的相关法律法规、政策、意见、标准规范、新闻报道等进行描述性统计分析和文本分析。本文采用的数据搜集方法如下:①从"北大法宝"-《中国法律检索系统》、统计年鉴及省级地方政府门户网站与省级政府数据开放平台上,以"数据开放""数据共享""数据开放共享"为检索词进行全文检索获取相关政府数据开放的政策,主要形式为《xx 市公共数据开放管理办法(草案)》《xx 省政务数据资源管理暂行办法》《xx 省公共数据和电子政务管理办法》等;②运用搜索引擎,使用"政府数据""开放""共享""标准""规划"等检索相关法律法规与政策、省级地方政府数据开放年度工作计划、领导讲话支持的新闻报导等相关内容。数据收集截至时间为 2019 年 12 月。表 1 为法律法规与政策、意见及标准规范等样本数据来源(部分)。需要说明的是,有的省级政府目前已经制定了专门的数据开放标准规范,但不在本文的统计时间内,如 2020 年 10 月,山东省公布了《公共数据开放技术规范》《公共数据开放元数据规范》等相关标准规范;浙江省于 2020 年 6 月公布《浙江省公共数据开放与安全管理暂行办法》等内容。

表1　样本数据来源(部分)

区域名称	数据来源	平台域名
贵州	《贵州省政务数据资源管理暂行办法》《贵州省大数据发展应用促进条例》《贵州省政府数据数据分类分级指南》《贵州省大数据安全保障条例》《贵州省政务数据资源管理暂行办法》	http://data.guizhou.gov.cn
上海	《上海市公共数据开放暂行办法》《上海市政府办公厅关于推进政府信息资源向社会开放利用工作的实施意见》《上海市政府数据资源向社会开放年度工作计划及重点领域》《上海市公共数据开放分级分类指南(试行)》	http://data.sh.gov.cn
天津	《天津市促进大数据发展应用条例》《天津市数据安全管理办法(暂行)》	http://data.tj.gov.cn
福建	《福建省政务信息系统整合共享实施方案》《福建省政务数据管理办法》《福建省电子政务建设和应用管理办法》	http://data.fujian.gov.cn
海南	《海南省大数据开发应用条例》《海南省公共信息资源安全使用管理办法》《海南省公共信息资源管理办法》《海南省政务信息整合共享专项行动实施方案》《海南省促进大数据发展实施方案》	http://data.hainan.gov.cn
浙江	《浙江省公共数据和电子商务管理办法》《浙江省人民政府办公厅关于印发2018年浙江省政务公开工作要点的通知》《浙江省人民政府关于印发浙江省促进大数据发展实施计划的通知》	http://data.zj.gov.cn
北京	《北京市政务信息资源管理办法(试行)》《关于全面推进政务公开工作的实施意见》《北京市交通出行数据开放管理办法(试行)》	http://data.beijing.gov.cn
山东	《山东省电子政务和政务数据管理办法》《山东省数字政府建设实施方案(2019-2022年)》《山东省人民政府办公厅关于印发山东省政务信息系统整合共享实施方案的通知》《开展基层政务公开标准化规范化试点工作实施方案》《山东省加快推进"互联网+政务服务"工作方案》《山东省人民政府关于促进大数据发展的意见》	http://data.sd.gov.cn
广东	《广东省数据开放和共享系列标准》《广东省政务信息系统整合共享实施方案》《广东省人民政府关于印发广东省促进大数据发展行动计划(2016-2020年)的通知》	http://gddata.gd.gov.cn
河南	《河南省政务信息系统整合共享实施方案》《河南省加快推进"互联网+政务服务"工作方案》	http://data.hnzwfw.gov.cn
陕西	《陕西省政务信息系统整合共享实施方案》《陕西省政务信息资源共享管理办法》	http://www.sndata.gov.cn
宁夏	《宁夏回族自治区开展基层政务公开标准化规范化试点工作实施方案》《宁夏回族自治区大数据产业发展促进条例(草案)》《宁夏回族自治区全面推进政务公开若干重点工作实施方案》《关于全面推进政务公开工作的实施意见》	http://ningxiadata.gov.cn
江西	《江西省大数据发展行动计划》《关于加快推进全省政务数据共享工作方案》《2018年度省级政务信息资源共享和开放责任清单》、《江西省政务信息系统整合共享实施方案》	http://data.jiangxi.gov.cn

1.3　研究变量确定

1.3.1　结果变量的确定

政府数据开放的准备度衡量是比较复杂的问题，《中国开放数林指数》和《中国地方政府数据开放报告》是我国首个专注于评估政府数据开放水平的专业指数和报告，笔者采用《2019中国开放数林指数》中的准备度指数反映政府数据开放的准备程度，将准备度指数作为结果变量，对其进行降序排序，按中位数赋值，大于等于中位数的赋值为1，中位数以下的赋值为0。具体数据如表2所示：

表2　13个省级2019年政府数据开放准备度赋值结果

编号	区域名称	准备度	赋值
1	贵州	19.13	1
2	上海	18.65	1
3	天津	11.78	1
4	福建	11.60	1
5	海南	11.05	1
6	浙江	10.53	1
7	北京	10.18	1
8	山东	7.95	0
9	广东	6.75	0
10	河南	4.20	0
11	陕西	4.00	0
12	宁夏	2.70	0
13	江西	1.80	0

资料来源:《2019中国开放数林指数》

1.3.2　条件变量的确定

条件变量是QCA方法中的关键环节，变量的确定要遵循一定的理论和方法。国内政府数据开放准备度的相关研究较少，笔者以中国知网(CNKI)、万方、维普数据库为文献资源，以政府数据开放作为主题，与"准备"组配进行主题、篇名、关键词等检索，检索时间截至2020年11月，在剔除不相关和重复论文后，得到相关文献27篇。国内关于政府数据开放准备度的研究还较为欠缺，笔者对政府数据开放准备度相关文献进行阅读、提炼，根据现有研究框架及运用文献归纳法来确定政府数据开放准备度的前因条件。如前文所述ODB的开放数据准备度评估与政策紧密相关，目前国内"中国开放数林指数"评估专家委员会，基于数据开放的基本原则，以跨界、多学科、第三方、中立的专业视角，立足我国国情与数据开放实践，构建的政府数据开放准备度分析框架，其关键要素有：文本效力、内容要求、部门职责、领导公开支持、公布年度工作计划、制定标准规范、标准规范的专门性。笔者基于此，对这个"典型框架"作适当调整，对变量进行归纳、筛选、整合及变量的重新概念化。对以上分类做出重新拆分，具体操作如下：

(1)政府数据开放准备度之标准规范是指为了在一

定的范围内获得最佳秩序，经协商一致制定并由公认机构批准，共同使用和重复使用的规范性文件。13 个省级以上的政府数据开放地方，仅有广东与山东两省制定了专门性的数据开放标准，贵州省涉及到了有关数据开放的内容，目前我国的政府数据开放尚处于发展阶段，在数据开放的标准规范方面还不够成熟，笔者将“制定标准规范”与“标准规范的专门性”统一为“标准规范”。

（2）相关的政策法规是政府数据开放共享的基础保障，从法律法规和政策文件的等级与专门性考察“文本效力”指标。同时，拆分原有的“文本内容”核心要素，将其拆解为政府数据管理、数据利用、安全保障。数据管理关系到开放数据的标准、治理、开放方式等内容，其重要性不言而喻。数据利用是从公民角度，满足公民对数据的知情、服务和再利用创造价值的需求，在政府数据开放的基础准备阶段，数据利用是重要影响因素。数据开放的安全保障是开放的前提，它的作用贯穿政府数据开放的始终。数据管理、数据利用、安全保障是数据开放重要影响因素，因此，笔者将它们分开独立成 3 个条件变量。

（3）政府数据开放准备度的组织与执行是指政府数据开放需要有效的管理体制、领导支持与工作计划，主要包括主管部门层级与相关部门职责、领导公开支持、是否公布年度工作计划。通过阅读政策文本与相关文献发现，部门职责包括政府数据开放主管部门及相关部门的职责分工，是保障数据开放工作的组织基础；领导（地方高层领导）的公开支持有利于开放数据工作顺利展开；制定并向社会公开数据开放年度工作计划既有助于政府部门自身明确阶段性目标与任务，也能给数据利用者明确预期，并接受社会监督。部门职责、领导公开支持、公布年度计划三大因素主要考察政府数据开放中组织及执行的准备程度。

通过分析、归纳整理，最终提炼确定 8 项影响地方政府数据开放准备度的影响因素，包括：文本效力、数据管理、数据利用、安全保障、部门职责、领导公开支持、公布年度工作计划、标准规范，这 8 项关键要素是 13 个省级地方政府数据开放准备度的核心内容，可以作为条件变量进一步探究其影响机制。

2 变量赋值

在条件变量与结果变量确定后，将样本数据进行人工筛选、归纳、分析，对其进行统一的变量选择与说明，表 3 为条件变量与结果变量设定说明。

表 3 条件变量与结果变量设定

变量内容		判断说明	赋值
文本效力	专门针对政府数据开放的法规和政策	如贵州省、天津市出台地方性法规	1
	无明确规定	法规、政策关于政府数据开放的内容较笼统	0
数据管理	明确提及数据管理的规定	“政务数据开放应按照无条件开放、依申请开放和依法不予开放三类进行管理……”	1
	未明确数据管理	没有提及优先开放领域、编制数据开放目录与清单、数据开放质量等有关内容	0
数据利用	明确开放数据利用规范	“市和区人民政府及其有关部门应当加强农业农村大数据应用，完善基础资源数据库，搭建资源要素台……”	1
	未提及或模糊的数据利用规定	未提及或模糊的数据利用相关事宜	0
安全保障	明确规定数据使用的安全保护及保障机制	“统筹构建云上贵州系统平台物理层的整体安全防护体系和认证体系……”	1
	无明确规定	未提及数据使用的安全保障相关事宜	0
部门职责	界定清晰	“市经济信息化部门负责指导协调、统筹推进公共数据开放、利用和相关产业发展……”	1
	界定模糊	未制定明确的部门职责	0
领导公开支持	有明晰的、公开的领导支持	“公共数据整合开放，既是集中管理、高效便捷应用的需要，也是用好公共数据、发展数字经济的需要……”	1
	内容笼统、不明确	内容笼统、不明确	0
公开年度工作计划	制定并公开切实有效的数据开放年度工作计划	“总体目标：……2019 年，计划累计开放公共数据 4000 项，重点在医疗、旅游、交通、信用等领域深入推进”	1
	未公开或无明确规定	无公开工作计划	0
标准规范	正式发布并实施详细的开放数据标准规范和操作指南	如广东省《电子政务数据资源开放数据技术规范》	1
	未制定或不具有专门性数据使用的标准规范	未制定、不具有专门性数据使用的标准规范	0
准备度指数	高准备度	中位数及以上	1
	低准备度	中位数以下	0

资料来源：各地区相关法律法规、政策、标准规范等整理

（1）文本效力主要从法律法规、政策文件的等级和专门性两个方面考察。从地方政府门户网站及地方政府数据开放平台发布的相关法律法规、政策文件来看，可将其法律法规、政策文件等级分为地方性法规、地方政府规章、规范性文件、方案意见，如目前贵州省、天津市、海南省通过政府数据开放的地方性法规，上海市出台的《上海市公共数据开放暂行办法》，广东省颁布相关的规范性文件、北京出台与交通出行有关的部门规范性文件。就法规与政策的制定是否专门为政府数据开放来说，截至2019年12月，13个省级地方尚未出台专门针对数据开放的地方性法规、规章或规范性文件，出台的法规与政策多为针对公共数据资源共享、开放、管理及大数据发展促进等内容，仅上海市制定了专门针对数据开放的地方政府规章，还发布《上海市公共数据开放分级分类指南》方案意见。笔者将13个省级地方现有的相关法规与政策文本进行梳理，将对相关法律法规与政策中至少有一节专门针对政府数据开放的或专门针对数据开放的法规、政策和方案意见等内容的赋值为1；否则赋值为0。

（2）数据管理主要是对开放数据资源进行管理，通过对13个地方有关政府数据开放的地方性法规、地方政府规章、规范性文件及方案意见中的内容进行分析，将包括开放数据范围、开放数据目录与清单、开放数据质量等数据管理内容的赋值为1，如在《贵州省政务数据资源管理暂行办法》第六章“数据开放”部分，“对政务数据按照无条件开放、依申请开放和依法不予开放三类进行管理”，并规定“编制数据开放负面清单”，还要加强“对开放数据进行动态更新管理，确保开放数据及时有效”；否则赋值为0。

（3）数据利用是数据开放的根本目的，通过利用开放的数据提高政府透明度、效率，鼓励社会主体参与创新。我国政府尚未出台全国层面的关于政府数据利用的法律法规，与数据利用相关的政策多为意见、规定、方案等导向性政策文本，在本文数据搜集时间内，仅有贵州、天津、海南以地方立法的形式明确政府数据利用规范，如《贵州省大数据发展应用促进条例》《天津市促进大数据发展应用条例》《海南省大数据开发应用条例》等。另外，省级地方政府开放数据开发公共服务，推进行业政府数据资源挖掘和融合，例如2019年11月北京出台的《交通出行数据开放管理办法（试行）》。此类内容赋值为1，否则赋值为0。

（4）安全保障主要包括两个方面，一方面加强对数据的使用、安全和隐私方面的法律和规则制定[16]；另一方面为保障数据开放的推进，建立的协调、人员、资金等保障机制[17]。对各地与数据开放有关的法律法规、政策、计划等内容进行统计分析，将明确含有安全保障内容的赋值为1，未提及数据使用安全保障相关事宜的赋值为0。

（5）部门职责是政府数据主管部门及相关部门的职责分工，政府数据开放工作需要专门的数据主管部门及专门的数据管理人员负责对开放数据的统筹规划和实施[18]。将13个地方在“数据开放”相关的法规和政策中对部门职责做出具体明确规定的赋值为1，如浙江省设有大数据发展管理局，并规定“负责编制公共数据资源目录和开放目录……推进落实各级各部门信息系统互联互通……”[19]；否则赋值为0。

（6）领导公开支持是指地方高层领导（如党委正副书记、省长、市长等）的支持有力推进政府数据开放，其讲话内容明确涉及数据开放的经验、问题、意义与要求，赋值为1，如2019年1月12日福建省委常委福建省人民政府常务副省长张志南在出席福建省公共信息资源统一开放平台和海丝卫星数据服务中心正式开通仪式时强调推动数据与互联网、大数据、人工智能同实体经济深度融合及利用[20]；有些讲话虽然涉及到了数据开放，但内容较为宽泛或仅涉及要在数据开放上取得新进展的内容，赋值为0。

（7）公布年度工作计划是为政府部门自身明确阶段性目标，同时给数据利用者明确预期，接受社会监督而制定的公开年度数据开放工作计划。如上海市自2014年开始每年制定和公布《政府数据资源向社会开放年度工作计划及重点领域》。将明确向社会公众公布的年度工作计划赋值为1，否则赋值为0。

（8）标准规范的筛选标准为已经正式发布并实施详细的开放数据标准规范和操作指南，行业性的标准规范未纳入研究范围。如山东省针对各级政务部门的数据开放工作制定了公共数据开放的《基本要求》《数据脱敏指南》《开放评价指标体系》[21]；广东省“数据开放和共享系列标准”，具体包括《政务信息资源标识编码规范》（DB44/T2109－2018）、《电子政务数据资源开放数据技术规范》（DB44/T2110－2018）、《电子政务数据资源开放数据管理规范》（DB44/T2111－2018）3个地方标准[22]。对已正式发布并实施的、专门针对政府数据开放制定的标准规范赋值为1，否则赋值为0。

3 QCA 单变量必要性分析和条件组合分析

3.1 单变量必要性分析

QCA 分析的单变量必要性主要是通过计算一致性和覆盖率来判定,确定条件变量和结果变量之间是否存在充分性和必要性的关系。政府数据开放准备度影响要素单个变量对于结果变量的必要性分析,结果见表 4。一致性指标(consistency)用于判断条件(或条件组合)是否是结果变量的必要条件或充分条件,必要条件的计算公式如下:

$$Consistency(Yi \leqslant Xi) = \sum min(Xi, Yi) / \sum(Yi)$$

公式(1)

一般情况下,当 Consistency 指标大于 0.8 的条件,可以认为是结果的必要条件。当一致性满足后,再进一步分析覆盖率(coverage)指标,通过计算条件变量对于结果变量的交集,测量某个条件能解释结果的估计范围,该指标用来描述条件(或条件组合)X 对结果 Y 的解释力,覆盖率指标越大说明 X 对 Y 的解释力越大。

表 4 单变量必要性分析

条件变量	outcome = 1		~ outcome = 1	
	consistency	coverage	consistency	coverage
文本效力	0.57	1.00	0.00	0.00
~文本效力	0.43	0.33	1.00	0.67
数据管理	1.00	0.64	0.67	0.36
~数据管理	0.00	0.00	0.33	1.00
数据利用	0.86	0.67	0.50	0.33
~数据利用	0.14	0.25	0.50	0.75
安全保障	1.00	0.58	0.83	0.42
~安全保障	0.00	0.00	0.17	1.00
部门职责	0.71	0.63	0.50	0.38
~部门职责	0.29	0.40	0.50	0.60
领导公开支持	0.57	0.80	0.17	0.20
~领导公开支持	0.43	0.36	0.83	0.63
公布年度工作计划	0.14	1.00	0.00	0.00
~公布年度工作计划	0.86	0.50	1.00	0.50
标准规范	0.14	0.25	0.50	0.75
~标准规范	0.86	0.67	0.50	0.33

资料来源:fsQCA3.0 软件生成

在必要性检验中,当结果变量为"outcome = 1"的一致性检验时,即在政府数据开放准备度高的结果下,数据管理、数据利用、安全保障这 3 个条件变量取值大于 0.8,即一致性大于 0.8,其余条件变量的一致性低于 0.8,说明在单个变量中数据管理、数据利用、安全保障是结果变量的必要条件。在覆盖度的检验中,数据管理、数据利用、安全保障的覆盖率分别为 0.64、0.67、0.58,均未大于 0.8,解释力不太强。

在结果变量为" ~ outcome = 1"的一致性检验中,数据管理、数据利用的一致性小于 0.8,安全保障的一致性大于 0.8,这表明数据管理、数据利用这两种单项前因条件变量作为必要条件是不充分的,而安全保障作为必要条件是充分的。

3.2 条件组合的一致性和覆盖度分析

通过 QCA 分析软件对真值表进行布尔最小化运算,结果以低、高准备度为导向,得出 3 个解,分别为:复杂解(complex solution)、简约解(parsimonious solution)和中间解(intermediate solution)。简约解得到的结论比较简单,其启示性一般较差;中间解是较能说明问题、具有代表性的解,不会一成不变的按照变量设定而出现,又不会违背案例的事实,得到结论的启示性和普适性较好。笔者拟采用表 5 解释说明。

从表 5 看出,结果变量为"1"的 5 个组合的一致性和所有组合的总体一致性值均为"1",所有组合的总体覆盖度是 0.86,A1、A2、A3、A4 原覆盖度为 0.29,(这 4 种组合的解释力具有相似性),即为覆盖比例最高的 4 种组合,也是政府数据开放准备度高的典型组合:①文本效力 * 数据管理 * 数据利用 * 安全保障 * 部门职责 * ~公布年度工作计划 * ~标准规范;②数据管理 * 数据利用 * 安全保障 * 部门职责 * 领导公开支持 * ~公布年度工作计划 * ~标准规范;③文本效力 * 数据管理 * 数据利用 * 安全保障 * 部门职责 * 领导公开支持 * ~标准规范;④文本效力 * 数据管理 * 数据利用 * 安全保障 * 部门职责 * 领导公开支持 * ~公布年度工作计划。

4 研究结果及发展建议

通过前文对 A1、A2、A3、A4 四组的变量构成分析,可以发现政府数据开放准备度差异与条件变量的不同组合有着密切的关系,详细说明如下:

(1)在单变量必要性检验中,安全保障是保障政府数据开放准备度高必不可少的条件,它涉及政府数据安全与数据隐私保护等内容,主要依靠法律法规、政策、标准制定等对政府数据开放规范性要求。在 A1 ~ A5 组合中都存在,说明安全保障这一要素在促进政府数据开放准备度完善程度具有极其重要的作用。

(2)A1 ~ A4 组合中正面结果中显示数据管理、数

表 5 条件组合的一致性和覆盖度分析

组合	中间解条件组合	原覆盖度（raw coverage）	净覆盖度（unique coverage）	组合一致性（consistency）	总体覆盖度（solution coverage）	总体一致性（solution consistency）
A1	文本效力 * 数据管理 * 数据利用 * 安全保障 * 部门职责 * ～公布年度工作计划 * ～标准规范	0.29	0.14	1		
A2	数据管理 * 数据利用 * 安全保障 * 部门职责 * 领导公开支持 * ～公布年度工作计划 * ～标准规范	0.29	0.14	1		
A3	文本效力 * 数据管理 * 数据利用 * 安全保障 * 部门职责 * 领导公开支持 * ～标准规范	0.29	0.14	1	0.86	1
A4	文本效力 * 数据管理 * 数据利用 * 安全保障 * 部门职责 * 领导公开支持 * ～公布年度工作计划	0.29	0.14	1		
A5	～文本效力 * 数据管理 * 数据利用 * 安全保障 * ～部门职责 * ～领导公开支持 * ～公布年度工作计划 * ～标准规范	0.14	0.14	1		

资料来源：fsQCA3.0 软件生成

据利用、安全保障、部门职责这 4 种要素都是同时出现，表明这 4 个条件变量都很重要，即数据管理、数据利用、安全保障、部门职责需要同时具备，共同促进政府数据开放准备工作。

（3）对比研究结果可以发现，文本效力与领导公开支持存在交替使用的情况，也有同时存在的情况，因此文本效力与领导公开支持对结果变量没有显著影响，它们在不同的条件组合中以辅助条件的形式与数据管理、数据利用、安全保障、部门职责等要素共同作用产生不同的政府数据开放准备度结果。

（4）“公布年度计划”“标准规范”要素在 A1、A2、A3、A4 政府数据开放准备高的 4 种组合中以“～公布年度计划”（缺失公布年度计划）“～标准规范”（缺失标准规范）存在的有 3 种组合，说明“公布年度计划”“标准规范”可以在政府数据开放准备度高组合中缺失，QCA 研究结果显示，这两种条件因素在政府数据开放准备度中未具备显著的促进作用。

依据上述研究结论，结合我国政府数据开放实践，笔者给出如下建议：

（1）加强数据管理、数据利用、安全保障等要素对政府数据开放准备度的影响。政府数据开放很大程度上是由供应带动的数据管理贯穿于数据流动的每个阶段，它保障了数据流正常、高效率运转，涉及开放数据范围、目录清单编制、开放数据质量管理、开放方式要求等内容[23]。数据利用应制定相应的法律法规与政策，消除企业、个人对数据采集和开放利用的顾虑，更好地实现数据开放与利用，鼓励各地区采取不同的方式，如运用竞赛的形式宣传推广，或者搭建“众创空间”数据应用创新服务平台[24]。开放政府数据涉及国家秘密、个人隐私、商业机密等问题，安全保障是数据开放长远、可持续发展的重要基础[25]，不仅要加强对数据的使用、安全和隐私方面的法律法规制定，还要建立如人员能力、资金保障等保障机制。国内外经验表明，清晰明确的政府数据开放法规与政策是推进政府数据开放的法治基础和重要依据，数据管理、数据利用、安全保障既是政府数据开放法规政策的重要内容，也是具体内容。目前，国家－省级－地级的开放数据资源政策构建及实施已经初步形成，出台一系列政策和法律，以“主动”的激励方式为主，如《上海市政府办公厅关于推进政府信息资源向社会开放利用工作的实施意见》[26]《浙江省公共数据开放与安全管理暂行办法》[27]。但相关的内容较为笼统，仅有如天津市等部分地方明确要求编制开放数据目录与清单，或仅有如贵州省等部分地区明确了数据开放前的数据审查。因此，在今后制定政府数据开放法规政策时应加强数据管理、数据利用、安全保障 3 方面内容的详尽性及明确性。

（2）注重条件要素组合的作用。通过上述分析得出结论，文本效力、领导公开支持、部门职责与数据管理、数据利用、安全保障等要素紧密相关、相辅相成。比如，文本效力是法律法规和政策文件的约束力，它与数据管理、数据利用、安全保障的内容是紧密联系，相辅相成的，如贵州省的《贵州省政务数据资源管理暂行办法》[28]、天津市的《天津市促进大数据发展应用条例》[29]，通过制定相对的法律法规与政策，对开放数据治理、开放方式、推进利用、安全保护、保障机制等方面有明确的规定和要求。领导公开支持与部门职责关系到政府数据开放的组织和执行，政府数据开放工作也离不开有效的管理体制，与数据管理、数据利用及安全保障同样存在紧密联系。因此，在地方政府数据开放准备工作中要注意各条件要素的组合作用，提升数据开放准备工作开展的效果。

（3）探索“公布年度计划”“标准规范”等要素的积极作用。仅从数据分析结果来看，公布年度计划、标准规范要素，对政府数据开放准备度还不具备显著的促进作用。但从政府数据开放的实践来看，目前广东、山东两省制定了专门针对数据开放的地方标准，进一步规范化、制度化政府数据开放实践；上海市连续制定和公开了专门针对政府数据开放的年度工作计划，如 2020 年上海市公布《上海市公共数据资源开放年度工作计划》[30]，从整体目标、重点任务、工作要求与保障措施 3 个方面，主动公开年度上海市公共数据开放的重点领域，特别是将“防疫复工”放在首位，保障新冠疫情后期的经济全面恢复。在政府数据开放准备基础整体较弱的情况下，公布年度计划、标准规范对各地方政府数据开放工作的影响尚不明晰，还需更进一步的探索。

5 研究结论

笔者将贵州、上海、天津、福建等 13 个省级地方政府数据开放准备度指数为结果变量，把省级地方政府数据开放准备度影响要素作为条件变量，运用定性比较分析方法（QCA）对条件变量的各种组合进行比较分析，发现政府数据开放准备度影响因素中，安全保障是必要条件，数据管理、数据利用、部门职责、安全保障 4 种要素需同时具备，文本效力、领导公开支持以辅助条件形式与数据管理、数据利用、安全保障、部门职责要素组合在一起使用，而公布年度计划、标准规范要素在政府数据开放准备度中未具备显著的促进作用，未来还要加强探索。

不可否认，csQCA 分析只能处理二分类变量，对政府数据开放准备度的条件变量和结果变量必须被校准为 0 或 1，赋值为 1 表示属于该集合，0 表示不属于该集合，在处理相关案例类别过程中，不可避免地会存在一定的主观性。此外，笔者着重选择省级地方政府，后续研究可以扩大研究样本范围，考虑将数据开放准备度表现突出的地级城市（如贵阳市、济南市、成都市等）纳入研究，以调整和深化研究结论对政府数据开放实践的指导作用。

参考文献：

[1] United Nations. Guidelines on open government data for citizen engagement [EB/OL]. [2020 - 06 - 15]. http://www.gaportal.org/resources/detail/guidelines-on-open-government-data-for-citizen-engagement.

[2] 国务院关于印发促进大数据发展行动纲要的通知[EB/OL]. [2020 - 06 - 15]. http://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm.

[3] 政府工作报告——2019 年 3 月 5 日在第十三届全国人民代表大会第二次会议上[EB/OL]. [2020 - 07 - 03]. http://www.gov.cn/premier/2019-03/16/content_5374314.htm.

[4] Open Data Index. [EB/OL]. [2020 - 07 - 01]. https://blog.okfn.org/tag/global-open-data-index/.

[5] DAVIES T. Open data barometer-2013 global report[R/OL]. World Wide Web Foundation. [2020 - 07 - 15]. http://www.opendataresearch.org/dl/odb2013/Open-DataBarometer-2013-Global-Report.pdf.

[6] 张晓娟，孙成，向锦鹏. 基于开放数据晴雨表的我国政府数据开放提升路径分析[J]. 图书情报知识，2017(6)：60 - 72.

[7] World Bank's Open Government Data Working Group. Part a: open data readiness assessment users' guide [EB/OL]. [2020 - 07 - 20]. http://opendatatoolkit.org/docs/odra/.

[8] 2019 年中国开放树林指数[EB/OL]. [2020 - 07 - 19]. http://ifopendata.fudan.edu.cn/.

[9] RAGIN C. Configurational comparative methods qualitative comparative analysis (QCA) and related techniques[M]. New Delhi: India Safe Publications，2008.

[10] ISHIDA A, YONETANI M, KOSAKA K. Determinants of linguistic human rights movements: an analysis of multiple causation of LHRs movements using a Boolean approach[J]. Social forces, 2006(4)：1937 - 1955.

[11] 程聪，贾良定. 我国企业跨国并购驱动机制研究——基于清晰集的定性比较分析[J]. 南开管理评论，2016，19(6)：113 - 121.

[12] 伯努瓦里 豪克斯，查尔斯 C 拉. QCA 设计原理与应用——超越定性与定量研究的新方法[M]. 杜运周，李永发，译. 北京：机械工业出版社，2017.

[13] 邓胜利，付少雄. 定性比较分析（QCA）在图书情报学中的应用——以网络社区健康信息搜寻影响因素研究为例[J]. 情报理论与实践，2017，40(12)：23 - 28，11.

[14] 杨金龙，胡广伟. 移动学习采纳转化为持续的动因及其组态效应研究[J]. 情报科学，2019，37(7)：125 - 132.

[15] 杨洋洋，谢雪梅. 基于 QCA 的网络舆情热度影响因素构型分析[J]. 情报杂志，2019，38(5)：142 - 146.

[16] 洪京一. 从 G8 开放数据宪章看国外开放政府数据的新进展[J]. 世界电信，2014(Z1)：55 - 60.

[17] 联合国经济和社会事务部. 2016 联合国电子政务调查报告——电子政务促进可持续发展[R/OL]. [2020 - 05 - 04]. http://workspace. unpan. org/sites/Internet/Documents/UNPAN96420.pdf.

[18] 刘新萍，袁佳蕾，郑磊. 地方政府数据开放准备度研究：框架与发现[J]. 电子政务，2019(9)：2 - 11.

[19] 浙江省大数据发展管理局/机构职能[EB/OL]. [2020 - 05 - 15]. http://www.zj.gov.cn/col/col1629359/index.html.

[20] 加快数据汇聚开放共享推进大数据产业发展[EB/OL]. [2020

－05 － 15]. https://data. fujian. gov. cn/odweb/news/newsDetail. htm? news_id = 35.

[21] 山东省地方标准[EB/OL].[2020 － 06 － 15]. http://data. sd. gov. cn/odweb/news/newsDetail. htm? news_id = 51.

[22] 广东省数据开放和共享系列标准正式发布[EB/OL]. [2020 － 06 － 15]. https://gddata. gd. gov. cn/article/article/toArticleDetails/48.

[23] 张磊. 我国政府数据开放机制研究[J]. 情报探索,2015(9):25 －27.

[24] ZUIDERWIJK A, JANSSEN M. Open data policies, their implementation and impact: a framework for comparison[J]. Government information quarterly, 2014,31(1):17 －29.

[25] World Wide Web Foundation. Open data barometer-ODB global report. Third Edition[R/OL]. [2020 － 05 － 04]. http://opendatabarometer. org/3rdEdition/report/#readiness.

[26] 上海市人民政府办公厅印发《关于推进政府信息资源向社会开放利用工作实施意见》的通知[EB/OL]. [2020 －06 －08]. http://www. shanghai. gov. cn/nw2/nw2314/nw2319/nw22421/nw44400/u26aw39095. html.

[27] 浙江省公共数据开放与安全管理暂行办法[EB/OL].[2020 － 06 － 17]. http://www. zj. gov. cn/art/2020/6/17/art_1229017137_557682. html.

[28] 贵州省政务数据资源管理暂行办法[EB/OL].[2020 - 06 － 08]. http://www. guizhou. gov. cn/zwgk/zcfg/szfwj_8191/qfbf_8196/201709/t20170925_823937. html.

[29] 天津市促进大数据发展应用条例[EB/OL].[2020 －06 －08]. http://zscq. tj. gov. cn/zcfg/qt/201912/t20191203_147664. html.

[30] 上海市公共数据资源开放2020年度工作计划[EB/OL].[2020 －07 －08]. http://www. e-gov. org. cn/article-171842. html.

作者贡献说明：

朱玲玲：负责论文研究设计、数据收集、论文撰写及论文修改；
茆意宏：负责论文思路指导、论文修改；
朱永凤：数据收集与论文修改；
马坤坤：参与文献调研、论文修改。

Identification of Key Influencing Factors of Government Data Openness Readiness: A Case Study of Provincial Local Governments

Zhu Lingling[1] Mao Yihong[1] Zhu Yongfeng[2] Ma Kunkun[1]

[1] College of Information Management, Nanjing Agricultural University, Nanjing 210095

[2] College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211106

Abstract: [**Purpose/significance**] The readiness of government data opening is the perfection of the preparation for providing higher value data to the society, which is the foundation and support of open data. Based on the analysis of the influencing factors of local governments' data openness readiness, this paper proposes the path to promote the sustainable practice of data openness. [**Method/process**] Using qualitative comparative analysis method (QCA), the data openness readiness policies of 13 provincial governments are taken as the research samples. This paper analyzes the "combined effect" and "interactive relationship" of various factors affecting the readiness of government data opening. [**Result/conclusion**] The study found that security assurance is an indispensable prerequisite for high readiness for government data openness. The four elements of data management, data utilization, departmental responsibilities, and security assurance must be available at the same time. Text effectiveness and leadership public support are used in the form of auxiliary conditions combined with data management, data utilization, security assurance, and departmental responsibilities.

Keywords: open government data readiness QCA influence factors